“Deep Learning Based Assessment of Student Engagement in Classroom using Observable Actions and Emotions”

Ph.D. Synopsis
Submitted to
Swami Ramamand TeerthMarathwada University, Nanded

for
Doctor of Philosophy
in the subject
Computer Science
Under the Faculty of
Science and Technology

Submitted By
Mr. Shivling Shankarrao Patil
Research Student

Under the Guidance of
Dr. Nilesh Kailasrao Deshmukh
Associate Professor
School of Computational Sciences
Swami Ramanand Teerth Marathwada University, Nanded
2023-2024

Deep Learning Based Assessment of Student Engagement in Classroom using Observable Actions and Emotions

1. Introduction

In modern educational settings, understanding and enhancing student engagement is a pivotal factor in optimizing the learning experience and academic outcomes. The integration of deep learning techniques into education has paved the way for innovative approaches to assess and gauge student engagement in classrooms. This research delves into the captivating realm of "Deep Learning Based Assessment of Student Engagement in Classroom using Observable Actions and Emotions."

In traditional education, assessing student engagement has primarily relied on subjective measures, such as surveys and self-reporting, which can be limited by biases and inaccuracies. However, recent advancements in deep learning have unlocked the potential to objectively evaluate engagement by analyzing observable actions and emotions exhibited by students during learning activities. This research capitalizes on this technological progress to develop a robust framework that harnesses the power of deep learning algorithms to interpret a rich spectrum of student behaviors.

The proposed research acknowledges that engagement is a multifaceted construct encompassing not only overt actions like participation and attention but also subtle emotional cues like interest and confusion. By amalgamating deep learning methodologies with the intricacies of human expression, this study seeks to create a more nuanced and comprehensive approach to assessing student engagement. The utilization of advanced computer vision and affective computing techniques aims to extract valuable insights from students' facial expressions, body language, and interactions with educational materials.

Modern education systems rely heavily on learning environments in which students engage with a real-tutor (teacher) or auto-tutor to acquire the necessary information. Due to human-to-human connection, instructors may engage with pupils based on their observable actions, emotions, etc. in conventional learning. In this age of the smart city, there are several learning environments used in the teaching-learning process, including social e-learning, collaborative e-learning, m-learning, etc. Even there are smart campus that employs webinars and auto-tutors in addition to conventional learning and e-learning.

These different learning environments may be even more customised by including emotional and behavioural engagement analysis tools, which can automatically predict the students’ participation and helps modify the teaching style accordingly to improve the

teaching-learning process.Affective computing identifies, interprets, processes, and simulates human emotions [1]. Affective Computing may be used to assess students' involvement via their nonverbal indicators. It is a distinct field that encompasses the cognitive, emotional, and behavioural elements of people. The next sections give further information on the different learning environments utilised in the teaching-learning process, students' affective content analysis, automatically assessing students' involvement, and other relevant literature that leverages affective computing in the education domain.

This research holds substantial promise not only for educators striving to tailor their instructional methods but also for policymakers aiming to refine educational policies based on accurate engagement assessments. Furthermore, the findings from this study could open doors to the development of real-time engagement monitoring systems that provide timely feedback to both educators and students, fostering a dynamic and interactive learning environment.

2. Literature Review

2.1 Video Summarization

Automatic Summarization of Surveillance Videos ASOSV. Prior knowledge of the data, processing cost, and scalability always matter. Lu et al. developed unsupervised, efficient, scalable key-frame selection for surveillance footage summarising. TheMartingale Test is used for abnormality testing, and change detection per frame is based on RPM (Randomized Power Martingale), with the next value expected to be the same as the current value. According to the authors, key frame selection was subjective and result comparisons used older technology. Frame-based, frame-clustering, and subspace analysis Findings showed that extracting attributes took more time. Each frame was analysed for trustworthiness. The authors manually selected the group parameter's frame count[2].

Event-based summarization detects abnormal video occurrences. The system may summarise a video by watching and recognising changes, which are then submitted to a classifier to uncover abnormal traits [3]. A summarising algorithm combines the selected frames into a new set of usable frames. Event-based techniques are static algorithms that choose keyframes without detecting shot boundaries and eliminate data redundancy. Long-Term Identity Awareness of t racking Multiple People Face and non-negative matrix optimization are used in Video Summarization to observe individuals[4]. Frame entropy generates dynamic and static events [5]. Using the entropy of the frames to compare them aided event-based summarization. Video processing can classify occurrences. To aid in temporal data analysis, the categorization needs data labels[6].

2.2 Face Detection

Face detection is required for face identification, facial attribute classification, face editing, face tracking, etc [7, 8]. Reliable and effective field face identification remains a

difficulty despite progress in spontaneous face recognition in recent decades. This is due to variables such as body position, expression, scale, lighting, perspective and occlusion. However, face identification is more sensitive to changes in size than it is to variations in aspect ratio.

Early face recognition systems used hand-crafted image features and a classifier to find face regions. Haar Cascades classifier [9] and HOG followed by SVM [10] are two classic face detection methods. These works advance the state of the art. Face recognition on difficult pictures with many variables was limited. Researchers have developed a plethora ofmodel architectures in response to the recent success of deep learning in computer vision.

Using convolutional neural networks, Li et al. developed one of the first deep models for detecting human faces. Multiple phases of low resolution work to swiftly eliminate background areas, while the final high resolution phase thoroughly examines a shallow volume of permutations [11]. Face recognition and positioning were incorporated by Zhang et al. This study use three layers of deep convolutional neural networks to provide location predictions for faces and landmarks. In addition, they presented a web-based method for efficiently mining problematic samples [12]. Zhang et al., in their future work, combined CNN layers specifically dedicated to face and non-face classification. Contextual CNN architecture with two streams, which incorporates body part data to improve face recognition. They developed a data routing technique for training that allows for individual layers to get instruction from a variety of data sources, resulting in more capable models at the deepest layers[13]. Additional research uses cascaded CNN architectures after "CNN cascaded face-recognition training" [14] and "Face detection strategy based on cascaded convolutional networks"[15].

2.3 Deep learning for emotional analysis

Face gesture-based emotion detection systems provide the interpretation, labelling, and inference of cognitive affective states from a video picture of a face in real time. Researchers claimed that expressions of a face are triggered over a time span when an emotion is experienced, and that emotions can be interpreted by recognising the facial expression that is associated with it [16]. Face gesture-based emotion detection algorithms can analyse, classify, and infer cognitive affective states from a face video picture in real time. It is claimed that facial expressions are triggered throughout time when an emotion is felt, and that emotions may be identified by recognizing the facial expression associated with them [17, 18]

A survey of current advancements in the field of behavioural signal processing that places major attention on the advantages and limits of developing methodologies and technologies based on i) analysing human behaviours by altering multiple data modalities and (ii) real-world application [19]. Important are strategies for collecting emotional labels that demonstrate the similarities and differences across the categories of emotional labels. In this study, the authors have attempted to identify and analyse themany self-reported and perceivedmeanings of emotion[20]. A comparison examination of each approach’s accuracy,

manner of execution, benefits, and drawbacks illustrates discussions about facial expression methods.

2.4 Analysis on Student Engagement

There are just a few of systems that have researchedmultimodal impact detection in the classroom setting [21, 22]. Existing intelligent tutoring systems, auto-tutors, and humanoid robots judge students' level of interest by analysing facial expressions and other body parts linked to behavioural cues. Students' engagement in various studies are evaluated using learning-related emotions such as anger, boredom, confusion, disgust, curiosity, disgust, surprise, and frustrations[23]. For the examination of students' involvement with humanoid robot tutors, both positive and negative learning emotions, such as frustration, confusion, happiness, and despair, are included [24]. Whitehill et al. offered four behavioural engagement levels to measure the students' participation in a laboratory context [25].Behavioral patterns are subjective ad they were analyzed for student's participation in various classes based on two classes: Good or Bad. The method works for nonverbal behavior estimation and the study was carried out for an online class environment [26]. Analysis Students engagement can be established based on 4 classes (Engaged, Boredom, Frustrated and confused). Provided the network is able to consistently identify an emotion from the face [27].

2.5 Databases of Students' Affective States

In the vast majority of standard datasets, emotional states are classified according to the subjects' facial expressions. Ekman's four primary emotions serve as the basis for this classification system. Researchers have done little to discover how to properly identify and classify students' emotional states when they are in a classroom. There were a total of 29 male participants, and 21 female participants adding to a set of four hundred and twenty eight video clips in the Indian Spontaneous Expression (ISE) database thatwas published byHappy et al. They have divided the emotion into four basic states of joy, sadness, shock, and repugnance. Besides mild to strong expressions of emotion, both natural and performed,may be found in this repository of face images[28]. An E-learning database in connection with Ekman's emotions consisting of twelve female and 15 male participants was proposed by Riaz et al. [29].

Emotional involvement and behavior analysis dataset was provided by Whitehill et al. based on four classes for an E-learning platform. [30]. Several more research studies have been conducted on the e-learning system, but none of them use hand gestures or body postures for emotional state identification [31].

2.6 Major Gaps in the Literature Survey

To conclude, the significant gaps in the extant literature are as follows:

- The previously published research did not take into consideration any data gathered from large classes, webinars, or online courses.
- Emphasis on improving a deep network for a specific emotional understanding is not explored
- Many of the researches uses transfer learning or existing network with image sizes that are too huge.
- The existing algorithms for group level score prediction evaluate group level engagement predictions based on only textual data.
- Affective computing for the betterment of invariant, multi scale, emotional and behavioral analysis are not looked into.
- There hasn't been a lot of research done to ensure the accuracy of face detection in video.
- Although a number of publications havemade use of video summaries in relation to a certain topic, there is no one approach that can be used consistently to summarize Videos in relation to different subjects.

3. Problem Statement

The purpose of this study is to provide an unobtrusive approach of analysing students'emotional and behavioural involvement using nonverbal indicators that can be used in both e-learning and classroom settings. As a result, the research challenge is as follows: Investigations On Affective Computing to Improve Classroom Engagement Analysis in Higher Education by Deep learning.

4. Research objectives

The Research Objectives are

- To enumerate the available techniques in video Summarization and to develop an improvised keyframe identification algorithmusingHoG Features, adaptive thresholding and LocalMaxima estimation.
- To improve face detections in a video by reducing false detections and utilize hierarchical video summarization in a video.
- To improve the classification accuracy for emotions in a temporal environment through Progressive filter arrangement and shallow residual networks.
- To create a storyboard of behavioral analysis of students for a smart classroom based for Higher Education in India.

5. Research Methodology

The research methodology adopted for conducting a study titled "Deep Learning Based Assessment of Student Engagement in Classroom using Observable Actions and Emotions." The

methodology elaborates on the research design, data collection and preprocessing, model development, evaluation techniques, and ethical considerations.

a. Research Design:

The research design is primarily exploratory and analytical in nature, aiming to investigate the relationship between observable student actions, emotions, and engagement levels using deep learning techniques. The study will utilize both quantitative and qualitative data to ensure a comprehensive analysis of student engagement.

b. Data Collection:

i. **Observable Actions:** Video recordings of classroom sessions will be collected using multiple cameras strategically placed to capture students' facial expressions, body language, interaction with learning materials, and interactions with peers and instructors.
ii. **Emotion Data:** Facial expression recognition technology will be used to analyze students' emotional states during classroom sessions. This data will be obtained through video analysis software capable of detecting facial landmarks and emotions.
iii. **Engagement Data:** Engagement levels will be measured using sensors that capture students' movements, posture, and attention to the learning environment. This data can include variables such as gaze direction, body movement, and responsiveness to instructional cues.

c. Data Preprocessing:

i. **Video Segmentation:** Classroom video recordings will be segmented into smaller time intervals to facilitate processing. Each segment will correspond to a specific classroom activity or event.
ii. **Feature Extraction:** Features related to facial expressions, body movements, and interaction patterns will be extracted from the video data. These features will serve as inputs for the deep learning model.
iii. **Emotion Labeling:** Emotions detected from facial expressions will be labeled using established emotion categories such as happiness, surprise, sadness, etc.

d. Model Development:

i. **Deep Learning Architecture:** A convolutional neural network (CNN) and a recurrent neural network (RNN) will be combined to create a multimodal model capable of integrating information from observable actions and detected emotions.
ii. **Feature Fusion:** The CNN branch will process observable action data, while the RNN branch will process sequential emotion data. The outputs from both branches will be fused and fed into a fully connected layer for engagement prediction.

e. Evaluation Techniques:

i. **Cross-Validation:** The dataset will be divided into training, validation, and test sets using a cross-validation approach to ensure model robustness.
ii. **Performance Metrics:** The model's performance will be evaluated using metrics such as accuracy, precision, recall, F1-score, and confusion matrices. In addition, qualitative analysis of false positives/negatives will provide insights into the model's limitations.

f. Ethical Considerations:

i. **Informed Consent:** Prior informed consent will be obtained from students, parents, and teachers before recording classroom sessions. Data privacy and anonymity will be maintained during all stages of the study.
ii. **Data Security:** Recorded videos and personal data will be securely stored and accessible only to authorized researchers to prevent any data breaches.
iii. **Bias Mitigation:** Steps will be taken to ensure the training data is diverse and representative, minimizing potential biases in emotion and engagement recognition.

6. Limitations

The study may face limitations due to variations in classroom environments, student behaviors, and cultural differences, which can impact the generalizability of the findings.

In conclusion, this research methodology outlines the systematic approach to investigate the deep learning-based assessment of student engagement in the classroom using observable actions and emotions. By employing a multimodal approach and addressing ethical considerations, the study aims to contribute to the understanding of how technology can enhance educational assessments.

7. Significance of Study

The research topic holds significant importance in the field of education, technology, and psychology. This study aims to leverage advanced deep learning techniques to assess and understand student engagement in classrooms based on their observable actions and emotions. The significance of this study can be highlighted through the following points:

i. **Enhancing Teaching Practices**: By accurately assessing student engagement, educators can tailor their teaching methods to better suit the needs and preferences of their students. This can lead to more effective and engaging classroom experiences, resulting in improved learning outcomes.

ii. **Personalized Learning**: The insights gained from this research can contribute to the development of personalized learning strategies. Deep learning algorithms can identify patterns in student behavior and emotions, allowing for the customization of educational content and approaches for individual students.

iii. **Early Intervention and Support**: Detecting signs of disengagement or lack of understanding in real-time can enable teachers to intervene promptly. This can prevent students from falling behind and provide the necessary support to ensure their academic success.

iv. **Objective Assessment**: Traditional methods of assessing student engagement often rely on subjective observations. Deep learning algorithms can provide an objective and data-driven approach to measuring engagement, reducing potential biases and inaccuracies.

v. **Insights into Learning Environment**: The study could reveal correlations between specific classroom activities, teaching techniques, and student engagement levels. This information can guide educators in refining their instructional strategies and classroom dynamics.

vi. **Technological Advancements in Education**: The integration of deep learning technology into the education sector showcases the potential of technology to revolutionize traditional teaching paradigms. This research aligns with the ongoing trend of incorporating AI and machine learning into educational practices.

vii. **Cross-Disciplinary Collaboration**: The research requires collaboration between education experts, psychologists, and technologists, fostering interdisciplinary partnerships that can lead to innovative solutions and a holistic understanding of student engagement.

viii. **Long-Term Implications for Students**: Improved engagement assessment can contribute to a more positive attitude towards learning, potentially reducing dropout rates and fostering a lifelong love for education.

ix. **Ethical Considerations**: As the study involves the analysis of student actions and emotions, ethical considerations regarding data privacy, informed consent, and responsible use of technology come into play. Addressing these concerns can set a precedent for ethical AI deployment in education.

x. **Research Contribution**: The study contributes to the academic and research community by expanding the knowledge base at the intersection of education and deep learning. It paves the way for future studies and advancements in understanding how technology can be harnessed to optimize educational experiences.

In summary, the significance of the study lies in its potential to transform the way educators perceive and address student engagement, offering insights that can lead to more

effective teaching practices, individualized learning experiences, and the responsible integration of advanced technologies in the classroom

8. Proposed Work Plan

Year 1: Foundation and Data Collection

Quarter 1-2: Literature Review and Problem Definition

i. Conduct an extensive literature review on deep learning, student engagement assessment, observable actions, and emotion recognition.
ii. Define the research problem and establish the significance of assessing student engagement through observable actions and emotions.

Quarter 3-4: Data Collection and Preprocessing

i. Identify suitable educational institutions and obtain necessary permissions for data collection.
ii. Develop or curate a diverse dataset comprising of video recordings of classrooms, encompassing various student engagement levels, actions, and emotions.
iii. Preprocess and annotate the dataset, labeling observable actions and emotions during different segments of the videos.

Year 2: Model Development and Validation

Quarter 1-2: Feature Engineering and Representation Learning

i. Explore various methods for feature extraction from video data, including visual and audio cues.
ii. Implement and experiment with techniques like convolutional neural networks (CNNs), recurrent neural networks (RNNs), and attention mechanisms to capture temporal dependencies in actions and emotions.

Quarter 3-4: Deep Learning Model Development

i. Design a deep learning architecture that integrates the extracted features for assessing student engagement.
ii. Experiment with different model architectures, loss functions, and hyperparameters to optimize performance.
iii. Implement transfer learning techniques to leverage pre-trained models for feature extraction.

Year 3: Evaluation, Analysis, and Application

Quarter 1-2: Model Training and Validation

i. Train the developed model on the preprocessed dataset, fine-tuning as required.
ii. Employ cross-validation and validation metrics to ensure the model's effectiveness in assessing student engagement accurately.

Quarter 3-4: Evaluation and Analysis

i. Conduct extensive evaluations of the trained model on new data from different classrooms and institutions.
ii. Analyze the model's performance across diverse demographic and cultural contexts to ensure robustness.
iii. Compare the model's results with existing assessment methods to demonstrate its superiority.

Quarter 1-2: Application and User Interface

i. Develop a user-friendly interface or application that educators can use to input classroom videos for engagement assessment.
ii. Integrate real-time analysis capabilities, if feasible, to provide immediate feedback to educators.

Quarter 3-4: Documentation and Dissemination

i. Prepare detailed documentation of the research methodology, data preprocessing, model architecture, and evaluation results.
ii. Write research papers for submission to reputable conferences and journals in the field of education technology, deep learning, and artificial intelligence.

Quarter 1-2: Extension and Future Work

i. Identify potential extensions of the research, such as assessing engagement in online classrooms or in different learning environments.
ii. Consider the ethical implications of implementing such a system and propose strategies to mitigate potential risks.

Quarter 3-4: Finalization and Presentation

i. Finalize research papers and submit them to conferences and journals.
ii. Prepare and deliver presentations at relevant conferences, workshops, and seminars.
iii. Share the findings, codebase, and data (if possible and ethical) with the academic community to encourage collaboration and further research in this area.

By following this proposed research work plan, you will lay the foundation for an in-depth investigation into deep learning-based assessment of student engagement using observable actions and emotions, contributing valuable insights to the field of education technology.

References

[1] Rosalind W Picard. Affective computing mit press. *Cambridge,Massachsusetts*, page 2, 1997.

[2] Guoliang Lu, Yiqi Zhou, Xueyong Li, and Peng Yan. Unsupervised, efficient and scalable key-frame selection for automatic summarization of surveillance videos. Multimedia Tools and Applications, 76(5):6309–6331, 2017.

[3] Xinhui Song, Li Sun, Jie Lei, Dapeng Tao, Guanhong Yuan, and Mingli Song. Event-based large scale surveillance video summarization. Neurocomputing, 187:66–74, 2016.

[4] Shoou-I Yu, Yi Yang, Xuanchong Li, and Alexander G Hauptmann. Long-term identity-aware multi-person tracking for surveillance video summarization. arXiv preprint arXiv:1604.07468, 2016.

[5] Uros Damnjanovic, Virginia Fernandez, Ebroul Izquierdo, and JoséMaríaMartinez. Event detection and clustering for surveillance video summarization. In 2008 Ninth International Workshop on Image Analysis for Multimedia Interactive Services, pages 63–66. IEEE, 2008.

[6] Padmashree Desai, Jagadeesh Pujari, C Sujatha, et al. Impact of multi-feature extraction on image retrieval and classification using machine learning technique. SN Computer Science, 2(3):1–9, 2021.

[7] MeiWang andWeihong Deng. Deep face recognition: A survey. Neurocomputing, 429:215–244, 2021.

[8] ShervinMinaee, Amirali Abdolrashidi,Hang Su,Mohammed Bennamoun, and David Zhang. Biometrics recognition using deep learning: A survey. arXiv preprint arXiv:1912.00271, 2019.

[9] Paul Viola andMichael Jones. Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001, volume 1, pages I–I. Ieee, 2001.

[10] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), volume 1, pages 886–893. Ieee, 2005.

[11] Haoxiang Li, Zhe Lin, Xiaohui Shen, Jonathan Brandt, and Gang Hua. A convolutional neural network cascade for face detection. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 5325–5334, 2015.

[12] Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, and Yu Qiao. Joint face detection and alignment using multitask cascaded convolutional networks. IEEE signal processing letters, 23(10):1499–1503, 2016.

[13] Kaipeng Zhang, Zhanpeng Zhang, HaoWang, Zhifeng Li, Yu Qiao, andWei Liu. Detecting faces using inside cascaded contextual cnn. In Proceedings of the IEEE International Conference on Computer Vision, pages 3171–3179, 2017.

[14] Hongwei Qin, Junjie Yan, Xiu Li, and Xiaolin Hu. Joint training of cascaded cnn for face detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 3456–3465, 2016.

[15] Shervin Minaee, Ping Luo, Zhe Lin, and Kevin Bowyer. Going deeper into face detection: A survey. arXiv preprint arXiv:2103.14983, 2021.

[16] Hao Tang, Wei Wang, Songsong Wu, Xinya Chen, Dan Xu, Nicu Sebe, and Yan Yan. Expression conditional gan for facial expression-to-expression translation. In 2019 IEEE International Conference on Image Processing (ICIP), pages 4449–4453. IEEE, 2019.

[17] Heysem Kaya, Furkan Gürpınar, and Albert Ali Salah. Video-based emotion recognition in the wild using deep transfer learning and score fusion. Image and Vision Computing, 65:66–75, 2017.

[18] Xavier Alameda-Pineda, Elisa Ricci, andNicu Sebe. Multimodal behavior analysis in the wild: An introduction. In Multimodal behavior analysis in the wild, pages 1–8. Elsevier, 2019.

[19] Biqiao Zhang, Georg Essl, and Emily Mower Provost. Automatic recognition of self-reported and perceived emotion: Does joint modeling help? In Proceedings of the 18th ACM International Conference on Multimodal Interaction, pages 217–224, 2016.

[20] Husam Salih and Lalit Kulkarni. Study of video based facial expression and emotions recognition methods. In 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), pages 692–696. IEEE, 2017.

[21] Athanasios Psaltis, Konstantinos C Apostolakis, Kosmas Dimitropoulos, and Petros Daras.Multimodal student engagement recognition in prosocial games. IEEE Transactions on Games, 10(3):292–303, 2017.

[22] Jorge Castellanos, Pablo A Haya, and Jaime Urquiza-Fuentes. A novel group engagement score for virtual learning environments. IEEE Transactions on Learning Technologies, 10(3):306–317, 2016.

[23] K Dmello Sidney, Scotty D Craig, Barry Gholson, Stan Franklin, Rosalind Picard, and Arthur CGraesser. Integrating affect sensors in an intelligent tutoring system. In Affective Interactions: The Computer in the Affective LoopWorkshop at, pages 7–13, 2005.

[24] Amarjot Singh, Srikrishna Karanam, and Devinder Kumar. Constructive learning for human-robot interaction. IEEE Potentials, 32(4):13–19, 2013.

[25] Jacob Whitehill, Zewelanji Serpell, Yi-Ching Lin, Aysha Foster, and Javier R Movellan. The faces of engagement: Automatic recognition of student engagementfrom facial expressions. IEEE Transactions on Affective Computing, 5(1): 86–98, 2014.

[26] Mike Holmes, Annabel Latham, Keeley Crockett, and James D O'Shea. Near real-time comprehension classificationwith artificial neural networks: Decoding e-learner non-verbal behavior. IEEE Transactions on Learning Technologies, 11(1):5–12, 2017.

[27] Sidney D'Mello, Rosalind W Picard, and Arthur Graesser. Toward an affectsensitive autotutor. IEEE Intelligent Systems, 22(4):53–61, 2007.

[28] SL Happy, Priyadarshi Patnaik, Aurobinda Routray, and Rajlakshmi Guha. The indian spontaneous expression database for emotion recognition. IEEE Transactions on Affective Computing, 8(1):131–142, 2015.

[29] Sadia Riaz and ArifMushtaq. Emerging themes analysis of learner's aestheticemotions in e-learning environments. In 2016 Sixth international conference on information science and technology (ICIST), pages 399–405. IEEE, 2016.

[30] Jacob Whitehill, Zewelanji Serpell, Yi-Ching Lin, Aysha Foster, and Javier R Movellan. The faces of engagement: Automatic recognition of student engagementfrom facial expressions. IEEE Transactions on Affective Computing, 5(1): 86–98, 2014.

[31] TS Ashwin and RamMohana Reddy Guddeti. Affective database for e-learning and classroom environments using indian students' faces, hand gestures and body postures. Future Generation Computer Systems, 108:334–348, 2020.

Research Guide
Dr. Nilesh Kailasrao Deshmukh
Associate Professor
School of Computational Sciences,
Swami Ramanand Teerth Marathwada University,
Nanded

Research Student
Mr. Shivling Shankarrao Patil